Requested Patent: WO9833113A1

Title: VIRTUAL MEDIA LIBRARY ;

Abstracted Patent: WO9833113 ;

Publication Date: 1998-07-30 ;

Inventor(s): LONG ROBERT MYRON ;

Applicant(s): OVERLAND DATA INC (US) ;

Application Number: WO1998US01327 19980122 ;

Priority Number(s): US19970786702 19970123 ;

IPC Classification: G06F3/06 ;

Equivalents: AU5929398

ABSTRACT:

A media element library defines a virtual configuration different from the physical configuration of media and drives actually present in the library. A host computer system communicates with the library as if it were communicating with a conventional library having a physical configuration identical to the virtual configuration defined by the library. Library users configure the library to incorporate fault tolerance and/or performance enhancements such as data striping or mirroring, and the host computer utilizes the benefits of such enhancements by manipulating the virtual configuration in a conventional manner.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 3/06

A1

(11) International Publication Number: WO 98/33113

(43) International Publication Date: 30 July 1998 (30.07.98)

(21) International Application Number: PCT/US98/01327

(22) International Filing Date: 22 January 1998 (22.01.98)

(30) Priority Data:
08/786,702 23 January 1997 (23.01.97) US

(71) Applicant: OVERLAND DATA, INC. [US/US]; 8975 Balboa Avenue, San Diego, CA 92123-1599 (US).

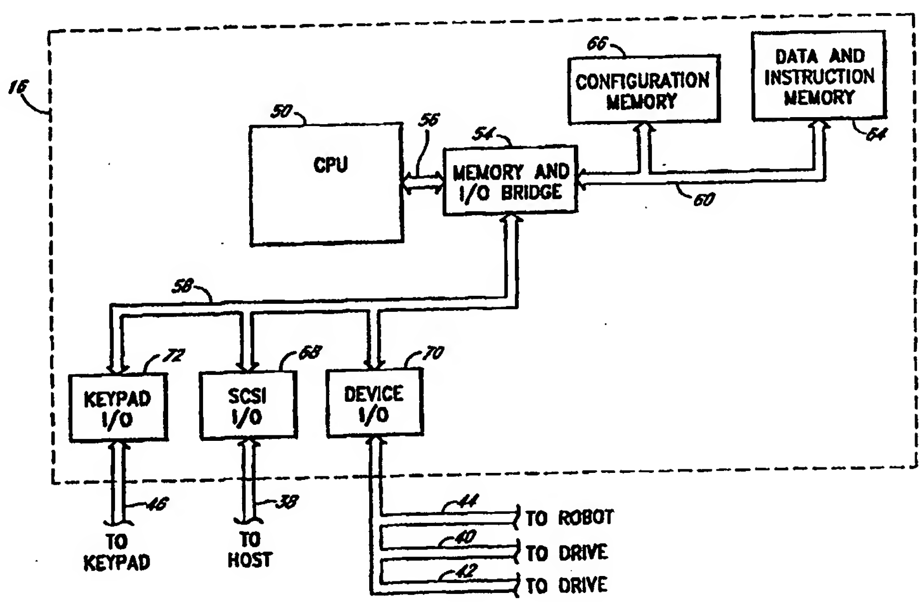
(72) Inventor: LONG, Robert, Myron; 615 Paseo Rio, Vista, CA 92083 (US).

(74) Agent: SIMPSON, Andrew, H.; Knobbe, Martens, Olson and Bear, 16th floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VIRTUAL MEDIA LIBRARY



(57) Abstract

A media element library defines a virtual configuration different from the physical configuration of media and drives actually present in the library. A host computer system communicates with the library as if it were communicating with a conventional library having a physical configuration identical to the virtual configuration defined by the library. Library users configure the library to incorporate fault tolerance and/or performance enhancements such as data striping or mirroring, and the host computer utilizes the benefits of such enhancements by manipulating the virtual configuration in a conventional manner.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

VIRTUAL MEDIA LIBRARY

Background of the Invention

The present invention relates to data processing systems and, more specifically, to automated data storage and retrieval systems which comprise a library of media elements as well as one or more drives for reading from and writing to the media elements in the library.

Magnetic tape cartridges, magnetic disks, and optical disks are all widely used as peripheral memory storage devices for computer systems. Large computer systems often operate in conjunction with external libraries having dozens of such media elements as well as the media element readers used to retrieve and record data. Although originally such media elements were selected and loaded manually, automated libraries were developed to expedite the handling of the media. These systems include means for accessing a desired media element, retrieving it from its storage position, and loading it into an appropriate reader. More recently, instruction sets have been created which define a communication protocol between the host computer system and the library. The instruction sets include commands to move media to different locations within the library and, of course, to load media into drives, place the media in a particular logical position, and read or write from and to specified regions of the media.

As data storage requirements for computer systems have increased from megabytes to gigabytes to terabytes, the development of automated media libraries has received considerable attention. Some embodiments of such libraries comprise a small number of media elements, six or ten being typical, and one or two drives housed in a single enclosure. Cabinet and even room sized systems have also been developed which hold a much larger number of media elements and drives, and which further comprise robotic arms, often translatable on all three axes, which remove media elements from storage and place them in drives. Furthermore, due to the increasing use of wide-area-networks, interconnected library systems forming a single dispersed database have become more common.

A second technique for digital data storage has also recently been developed, motivated by the desire to decrease the total storage cost per megabyte of a magnetic memory system without sacrificing performance or reliability. Such systems are commonly referred to as RAID systems (originally Redundant Array of Inexpensive Disks, now Redundant Array of Independent Disks), and they comprise an array of hard disk drives which are accessible in parallel by a host computer, thereby increasing the data rate above that available for each individual drive. Some data storage techniques for these arrays include the storage of multiple copies of data or error correction information to allow for the reconstruction of data files in the event of a failed drive. For example, data can be mirrored within the array. That is, each data block is stored on two separate drives. If one of the drives fails, the other is available for data reads. Another data management technique used with RAID systems is data striping. In these systems, data in a file is broken into multiple blocks, and the blocks are written sequentially to the drives in the array. This increases data rates during read and write operations because several drives are operating in parallel. Although no redundancy is directly achieved by striping, the addition of error correcting overhead such as the storage of parity information can provide for the reconstruction of a complete striped data file in the event of failure of any individual drive.

Some of the concepts of RAID data storage have been applied to magnetic tape arrays as well. Data striping in a tape array is described by Drapeau et al., "Striped Tape Arrays", 12th IEEE Symposium on Mass Storage Systems, 1993, pp. 257-265. However, the benefits of providing user configurable striped or mirrored tape libraries have not been appreciated prior to the present invention. Furthermore, although programs run by a host computer system can be designed which use the library instruction set to create mirrored or striped tape arrays within a tape library, this can be difficult to incorporate into host application software such as an automated data backup program, especially for existing application software not originally designed with mirrored or striped arrays in mind.

Accordingly, a need exists in the art which has been neither appreciated nor addressed prior to the present invention to provide a library of removable media which is easily configurable to incorporate striping, mirroring, and other RAID type data organizations, and which successfully interfaces with existing host application software.

Summary of the Invention

A library of removable media according to a preferred embodiment of the present invention comprises a library which reports to the host computer system a different media and drive configuration than is physically present in library. The "virtual" media configuration reported by the library to the host is determined by the library operating mode chosen by a user of the system. If the library is configured to mirror all incoming data, for example, the virtual media configuration reported to the host will include one half as many media elements (such as tape cartridges or optical disks) as are physically present in the library. In this case, when the host issues a write command, the library writes the data to two separate media elements. The write operation to the second media element is completely transparent to the host application program. This allows efficient communication between the host and the library, while creating increased flexibility in the data organization within the library itself.

Accordingly, one aspect of the present invention includes a method of communicating information between a host computer system and a data storage library. According to a preferred embodiment of this inventive method, a host computer sends a request for information regarding the number of media element storage locations in the library, and the library sends back a response to the request which includes a number of media element storage locations different from the number of media element storage locations physically present in the library. Thus, a "virtual library" is reported back to the host, and the host interfaces with the virtual library without requiring any knowledge of the data organization, such as striped or mirrored, within the library itself.

In another aspect of the present invention, a novel media element library is provided which comprises a plurality of storage locations for removable media elements, at least one media element drive, at least one robot for transporting media elements to and from the storage locations and drive. In this library, the number of storage locations, drives, and robots defines a library configuration. The library further comprises an interface to a host computer, wherein the interface accepts requests for configuration information from the library, and the library responds to the requests with configuration information different from the configuration defined by the existing storage locations, drive, and robot. The media elements may advantageously be tape cartridges, optical disks, or other removable media element. In a preferred embodiment, the library includes a non-volatile memory which stores information which determines the configuration information contained in the library response.

The present invention also includes a method of communication between a library and a host computer system. The novel communication method includes communicating a plurality of bits which are in a format that is interpreted by the host computer system as defining a specific number of media element storage locations. The plurality of bits is configured so as to be interpreted by the host computer system as a specific number of media element storage locations and/or media element drives which is different from the number of media element storage locations and/or drives which are physically present in the library.

Brief Description of the Drawings

FIG. 1 is a cutaway perspective view of the physical layout of a magnetic tape library.

FIG. 2 is a block diagram of the interconnections between the components of the tape library of FIG. 1 and a host computer system.

FIG. 3 is a block diagram of the library controller of FIG. 2 according to a preferred embodiment of the present invention.

FIG. 4 is an illustration of an industry standard format for a library status request command issued by a host computer.

FIG. 5A is an illustration of an industry standard format for the header portion of a library response to the status request command of FIG. 4.

FIG. 5B is an illustration of an industry standard format for a data portion of a library response to the status request command of FIG. 4.

FIG. 6 is an illustration of an industry standard format for a move medium request issued by a host computer.

FIG. 7 is an illustration of an industry standard format for a write request issued by a host computer.

Detailed Description of the Preferred Embodiment

Preferred embodiments of the present invention will now be described with reference to the accompanying Figures, wherein like numerals refer to like elements throughout. The terminology used in the description presented herein is intended to be interpreted in its broadest reasonable manner, even though it is being utilized in conjunction with a detailed description of certain specific preferred embodiments of the present invention. This is further emphasized below with respect to some particular terms used herein. Any terminology intended to be interpreted by the reader in any restricted manner will be overtly and specifically defined as such in this specification.

A cutaway view of an example tape cartridge library system is illustrated in Figure 1. It will be appreciated, however, that as described briefly above, media libraries come in many forms, and include a wide array of media types such as optical disk as well as magnetic tape, and also include a variety of different numbers of drives, tapes, and robotic arms for media transportation. Such differences in media type and hardware configuration do not affect the applicability of the present invention to these diverse styles of library. Accordingly, the system described in conjunction with the following figures is exemplary only.

Referring now to Figure 1, a cutaway perspective view of a preferred magnetic tape library system 10 incorporating two tape drives 12 and 13 and a tape storage magazine 14 is illustrated. Further provided inside

library 10 is a power supply (not shown) as well as a library control module 16. The circuitry of the library control module 16 and its response to commands from an external host computer (not shown in Figure 1) are described in more detail below with reference to Figures 3-7.

The library 10 also preferably includes a media transport robot 18 which may be threadably secured to a drive shaft 20, which is in turn secured just above the floor of the library 10 in bearings on the front and rear panels. Lateral motion of the media transport robot 18 is produced by a motor 24 which is secured to the inside of the library 10 housing, and which has a rotor coupled to the threaded drive shaft 20 by a drive belt 26. In a library configured as shown in Figure 1, the media transport robot moves back and forth along the drive shaft 20 so as to be adjacent to tape drive 12, tape drive 13, or any of the storage slots in the tape storage magazine 14. Overall operation of the library 10 proceeds as follows. Upon the receipt of commands from the library control module 16, the media transport robot 18 is positioned in front of a particular tape cartridge 30 in the magazine 14. The media transport robot 18 pulls the tape cartridge 30 from the magazine 14 into its own housing, and then travels toward the rear of the library 10.

When adjacent to the appropriate drive 12 or 13, the media transport robot 18 pushes the tape cartridge 30 out of its housing and into the drive 12, 13. It can be appreciated that a variety of functions can be performed by the media transport robot 18 in addition to loading a tape cartridge into a drive. For example, a tape cartridge could be moved to a different empty storage location in the magazine 14. In one preferable library embodiment, a pass through slot 32 is provided so that the media transport robot 18 can eject tape cartridges out of the library 10. If desired, a second media transport device can be utilized to move such an expelled tape cartridge to another library. Other library configurations adapted for use with the present invention in addition to that shown in detail in Figure 1 have been developed. For example, library systems are described in U.S. Patent No. 5,285,333 to Barr, et al., U.S. Patent No. 5,388,946 to Baur, U.S. Patent No. 5,429,470 to Nicol et al., and U.S. Patent No. 5,277,534 to Anderson et al. being several examples. The disclosures of each of these patents are hereby incorporated by reference in their entireties. It should also be kept in mind that libraries of other media types such as optical disks are also contemplated for use in conjunction with the present invention.

As is set forth fully below with respect to the remaining figures, the drives 12, 13 and media transport robot 18 are under the direct control of the library control module 16. The library control module 16 further receives commands from a host computer system which direct it to produce, in the library 10, action desired by the host system. In some embodiments of the library 10, the library control module 16 can also interface with users through a physical keypad/display unit 32. These interconnections are illustrated schematically in Figure 2.

As can be seen with reference to Figure 2, a host computer system 36 interfaces to the library 10 via a communication bus 38, which may advantageously be of Small Computer System Interface (SCSI) format. The host computer system 36 may be a personal computer, a mainframe, a local area network, or any of a wide variety of data processing apparatus well known to utilize media libraries for data storage. The communication bus 38 connects to the library control module 16 inside the library 10 itself. As is further shown in Figure 2, the library

control module 16 is also preferably interconnected to a keypad/display unit 32, for communicating a relatively limited set of commands and status messages between the control module 16 and a library user.

The library control module also communicates with each drive 12, 13 in the library through additional communication busses 40, 42, as well as communicating with the media transport robot 18 through another communication bus 44. These internal communication busses 40, 42, 44 may be configured in any convenient fashion and remain compatible for use with the present invention. For example, they may comprise a common physical set of communication lines, such as a single SCSI bus with the drives and media transport robot daisy chained. This can be advantageous in that existing tape drives are commercially available which are adapted to utilize the SCSI protocol for data transfer. Alternatively, the tape drive 12, 13, robot 18, and library control module 16 interface circuitry can be designed to use independent busses as shown explicitly in Figure 2, or as a common bus of non-SCSI format. In this case, a library 10 of higher performance may be created because the interface protocol can be tailored to the particular needs of a library system. However, the tape drives used in the library 10 are then unlikely to be functional outside of the particular library 10 which is being used. It can be appreciated that other configurations, such as a common bus for the tape drives, and a separate bus for the media transport robot, are also possible. Also illustrated in Figure 2 are media storage locations 46a through 46j. In the illustrated example, storage locations 46c, e, and f are vacant, and the remainder house magnetic tape cartridges.

In normal operation, the library 10 receives commands from an application program such as an automated data backup program running on the host computer system 36. These commands are interpreted by the library control module 16, which then carries out the commands by appropriately controlling the actions of the media transfer robot 18 and the drives 12, 13. As mentioned above, the library control module 16 may also interface with library users through a keypad/display unit 32 via another communication link 46. Simple commands such as a reset command and certain status messages may be relayed via this interface.

One aspect of the present invention is directed to the configuration of the library control module 16 and its responses to commands received from the host computer system 36. As is explained in detail below, a preferred media library 10 in accordance with the present invention will interpret and execute identical commands from the host computer system 36 in different ways depending upon how the library user configures the library control module 16. For example, a host computer command to write a block of data to a particular location on a media element may in actuality cause the library to write that block of data to a physically different location, or to several different locations, depending on the library configuration. In an especially advantageous embodiment, this different command execution is transparent to the host computer 36. As the physical nature of the library is in this way hidden from the host computer system 36, it may be stated that the host 36 is interfacing with a "virtual library", which appears to have a particular configuration of tapes, drives, and data which is different from the actual physical configuration of tapes, drives, and data in the library 10.

The implementation of the virtual library is preferably done inside the library control module 16 and, accordingly, a preferred library control module 16 is illustrated schematically in Figure 3. As is illustrated in this Figure, the library control module 16 includes a microprocessor circuit 50, which may advantageously comprise an

Intel 386, 486, or Pentium (TM) processor chip. The microprocessor 50 interfaces with a memory and I/O bridge 54 via a host bus 56. The memory and I/O bridge 54 in turn interfaces with an I/O bus 58 and a memory bus 60, and functions to buffer data and command transfers to and from the microprocessor 50. The memory and I/O controller 54 also create a direct memory access (DMA) channel between the I/O devices and the library control module mass memory 64.

In a preferred embodiment of the present invention, the memory of the library control module 16 comprises two portions, a data and instruction memory portion 64, and a configuration memory portion 66. The data and instruction memory is preferably a dynamic RAM of four or more megabytes, and stores instructions processed by the microprocessor 50 as well as data being transferred between the library 10 and the host computer system 36. Of course, it will be appreciated that memory size is not important to the operating principles of the present invention. The library control module 16 also preferably comprises a configuration memory 66. The configuration memory is for storing information relating to the user defined "virtual" configuration of the library. In a preferred embodiment, the configuration memory 66 is non-volatile, comprising a battery backed up RAM, EEPROM, flash memory, or other non-volatile memory device. This allows the library 10 to retain its virtual configuration information when the library is powered down.

The I/O bus 58 of the library control module 16 is connected to the host computer 36 through an industry standard SCSI interface circuit 68. The SCSI I/O circuit coordinates command and data flow between the host computer 36 and the microprocessor 50 of the library control module 16. The I/O bus 58 of the library control module 16 is also connected to a device I/O circuit 70 for interfacing with the robot 18 and the drives 12, 13 inside the library, and further is connected to a keypad I/O circuit 72 for interfacing with the front panel keypad 32. The device I/O circuit 70 may also comprise a SCSI interface similar to the SCSI interface 68 to the host computer, or may implement communication with the library robot 18 and drives 12, 13 using a different protocol.

In operation, the library control module 16 receives commands and data from the host computer system 36. In one preferred embodiment of the present invention, the commands issued by the host computer 36 are among the set of industry standard SCSI commands developed for host-library communication. These commands are interpreted by a program being run by the microprocessor 50, which controls the operations of the robot 18 and drives 12, 13 in accordance with host commands. As mentioned above and explained in detail below with reference to Figures 4 through 7, it is one aspect of a preferred embodiment of the present invention that the commands received from the host computer 36 are interpreted differently depending on the information stored in the configuration memory 66. It will be appreciated by those skilled in the art that various computer architectures could be utilized to produce the configuration dependent interpretation and execution of commands from the host computer system 36. Accordingly, the data flow illustrated in Figure 3 is intended to be exemplary rather than limiting.

The function of the library control module 16 and its processor 50 is the interpretation and execution of commands received by the host computer system 36. For example, the host computer may instruct the library to move a media element from one storage location to another, or may instruct the library to read or write at a particular location on a media element. In preferred embodiments of the present invention, as well as on many prior

art library systems, these instructions are interpreted and carried out by a software program running on the library control module. In a preferred SCSI interface implementation of a library, as in the example described herein, such commands issued by the host computer system 36 are defined by a standard library command set defined by industry convention. Although commands to the library 10 from the host computer 36 may be of any format as long as the library 10 is designed and programmed to understand the commands being issued by the host 36, the industry standard SCSI host-library communication protocol is used herein as an example of one preferred embodiment of the present invention.

Referring now to Figure 4, several of the significant features of the SCSI command referred to as a "read element status" command are illustrated. This command is issued by the host computer 36 to the library 10, and in response, the library returns information to the host computer regarding the library configuration including the number of media elements, drives, and robots. This command is comprised of several important portions. First, an opcode field 80 assigned to this instruction, and second, a series of parameters for informing the library what status information the host computer wishes to receive. The parameters include (1) an element type field 82, which indicates what type of element, i.e. media elements, drives, and/or robots, should be reported on by the library, (2) an element starting address 84, indicating the lowest address of the specified element type which should be reported by the library. The command also includes additional parameters such as an allocation length field, specifying how much memory has been allocated by the host for data received from the library in response to the command.

The standard format response to a one particular read element status command requesting status information regarding the media element storage slots is illustrated in Figures 5a and 5b. The response comprises an eight byte header illustrated in Figure 5a, followed by several twenty-six byte storage element status pages, one for each storage slot being reported. Referring now to Figure 5a, the header comprises information concerning the first element address reported 88, the number of elements being reported 90, and the total number of bytes of information 92 of available status information regarding the media element storage locations.

Figure 5b illustrates the content of each returned status page containing information regarding the media element storage slots reported by the library 10. The information in the status page includes (1) an element type code 94, which in this instance would be the element type code assigned to media element storage slots, (2) the address of the storage slot 96 being reported in this particular status page, (3) an access bit 98 indicating whether or not the robot 18 is allowed access to this slot, (4) a full bit 100, indicating whether or not a media element such as a tape cartridge is contained in the slot, and (5) a source address field 102, indicating the address which last contained the media element currently stored in the slot being reported.

It can be appreciated that a conventional library will respond to the read element status command of Figure 4 with information concerning the physical configuration of the library. For example, a read element status command could request information concerning any one or all of the of drives, storage slots, or robots in a media library. As seen above, a conventional library response would include the total number and addresses of the elements requested, the presence or absence of media elements in storage slots, as well as other information about the library 10, and

this information would correspond directly to the physical configuration of the library 10. Other host commands which are issued to a library also fit this paradigm.

In Figure 6, for example, the format of the "move media" command is illustrated. This command includes an opcode field 104, and additionally includes (1) the address 106 of the robot to be used to perform the transfer, (2) the current address 106 of the media element to be moved, and (3) the destination address 110 that the media element is to be moved to. In a conventional library, each storage slot is assigned an address, and accordingly, this command causes the transfer of one media element from a first defined physical location to a second defined physical location. There is, therefore, a direct correspondence between command parameters such as storage slot addresses and the physical configuration of the library. Another example is provided by the write command, the format of which is illustrated in Figure 7. This command includes (1) an opcode field 112, (2) a logical unit number 114 identifying the drive which is to write the data to the currently loaded and positioned media element, and (3) an indication of the transfer length 116, which may be specified in either the number of bytes or the number of multi-byte data blocks. Once again, the logical unit number has a one to one correspondence with a physical drive in the library which is directed to write data onto the currently loaded media element.

In the preferred library of the present invention however, the correspondence between addressed elements and action taken by the library 10 can vary such that the host computer system 36 sends commands based on a virtual library configuration that is different from the physical configuration actually present. Such an interface to a virtual library can improve library performance characteristics as seen by the host computer system 36 without complicating the set-up, maintenance, or creation of application software utilizing the library as a data storage device. Library performance can be improved with the use of data storage techniques which have been used in disk based systems such as the striped and mirrored arrays described above. By making the data storage technique being utilized by the library 10 transparent to the host computer system 36, however, the library's response to host commands must be altered, and must further be altered in different ways depending on the particular data storage technique the library is using. The following sections set forth several preferred configurations, and explain how the host issued command interpretation will vary, with particular reference to the commands illustrated in Figures 4, 6, and 7. A media library in accordance with the present invention preferably can be user configured in any one of the modes described in the sections below.

Single Mode:

This library configuration is functionally identical to a conventional library. Storage slot addresses correspond to particular physical storage slots, and each drive and robot is assigned its own logical unit number and address. Accordingly, commands issued by the host 36 are interpreted as acting upon the physical media, robot, and drives to which those commands are addressed. No library performance enhancements result from this configuration.

Spanned Mode:

In a spanned system, groups of media elements are treated as a single media element. In other words, blocks of 2, 5, 10 or any other size collections of media elements are reported as single media elements. In one such preferred system, all of the media elements in the library are considered a single "spanned" media element. In this case, if the library 10 incorporates N media elements such as tape cartridges, the data storage capacity of the library is N times the capacity of each individual media element. In such a completely spanned system, described in detail below, the first media element is completely filled before any data is written to the next media element and so on until all of the media elements are filled. End of tape (EOT) status during a write operation is therefore not returned to the host computer system 36 until after the last media element is full.

The virtual configuration of a completely spanned system is a single high storage capacity media element, and would typically be implemented in a library with a single drive. Therefore, when a virtually spanned library received the "read element status" command of Figure 4, the response, illustrated in general in Figure 5a, would list a single media element storage location defined by a single address, rather than the N media element storage locations that are physically inside the library. Furthermore, if any of the N media elements are in the tape drive, the library will return a status of empty in the "full" field 100 for the single virtual storage slot.

Because of this virtual configuration communicated to the host, the "move media" command of Figure 6 must also be executed in a different manner from the conventional (or "single mode") system. Because all N media element storage slots share a common address as far as the host computer 36 is concerned, the move medium command will only reference this one storage slot address. In other words, the host moves a single virtual media element between the drive and a single virtual storage location, even though there are N physical media elements in the library. It can therefore be appreciated that the library must choose which of the N media elements to move, as this cannot be specified by the host computer system 36. In one suitable embodiment, if the move media command is intended to initiate a move of the spanned media from the drive to the single virtual storage slot, this will be interpreted as a command to move whichever of the N media elements is currently in the drive back to the storage slot it was removed from. If the status of the spanned media element is stored in the virtual slot, corresponding to the state wherein all N media elements are loaded into their respective storage slots, a move media command intended to initiate a move from the virtual slot to the drive may be interpreted as a command to move the first media element from its slot to the drive. Alternatively, data access times may be improved somewhat by interpreting such a move media command as a command to put into the drive the media element that was last removed from the drive, or alternatively the media element that contains the last data block.

Although the write command itself, illustrated in Figure 7, is interpreted the same way as during conventional operation, the locate command, which precedes the write command and positions the media at the block address where the write operation is to begin, is made more complex in spanned system. Because the block address specified in the locate command may not be on the media element currently in the drive, the library may be required to return the media element in the drive to its storage location and to place the correct media element, which includes the requested block addresses, into the drive for writing. A spanned library therefore must keep a record of what data block addresses are on which of the physical media elements.

A similar situation occurs for reads or writes on a spanned system which extend from one physical media element to another. In a conventional system, a read or write which hits the end of the media such as the end of the tape in a tape cartridge produces an error condition. In a spanned system, however, the host remains unaware of the physical partitioning of its single piece of virtual media, and accordingly should not be notified of an error condition if it issues a read or write command extending across physical media boundaries. In these cases, the library will be programmed to automatically store the last media element and install the next media element into the drive to complete the read or write operation requested by the host computer system.

Striped Mode:

The concept of striping data has been used in the context of fixed disk RAID systems for some time, and the essential characteristics of this data organization may be implemented in a media library as well. In a striped system, data being written to the library will be broken up into a series of much smaller chunks of data, and the chunks are written sequentially to the drives in the library. For example, if the library includes four drives, chunks 1, 5, 9, 13, and so on would be stored on the media in the first drive, chunks 2, 6, 10, 14 and so on would be stored on the media in the second drive, chunks 3, 7, 11, 15 and so on would be stored on the media in the third drive, and chunks 4, 8, 12, 16, and so on would be stored on the media in the fourth drive. The advantage of data striping is that the apparent read/write speed of the library is equal to the speed of an individual drive times the number of drives in the library. In the above example, the data transfer rate is four times that of a single drive because chunks 1-4 are written essentially simultaneously, as are chunks 2-8, 9-12, and so on. Specific striping methodologies may vary in their details with respect to the size of the individual data chunks, and the physical position of the data chunks on the media. One such striping scheme in which the media is magnetic tape is disclosed in U.S. Patent No. 5,502,811 to Ripberger, the disclosure of which is hereby incorporated by reference. It will be appreciated by those of skill in the art that differences in striped data organization are not significant to the implementation of a library in accordance with the present invention.

In accordance with a preferred embodiment of the present invention, the host computer 36 is unaware of the striping process, and is interfacing with a virtual drive which has a read write data rate four times faster than each individual one of the four physical drives. In this case, therefore, when the host computer 36 requests status information regarding the drives present in the library, the library will report only one drive, even though it will contain several which are written to and read from simultaneously. Furthermore, when reporting to the host computer regarding the status of the media storage slots, the library will not report the existence of all physical slots to the host, but will instead report that the library has the number of physical slots divided by the number of striping drives. In the four drive example described above, therefore, the storage slots for physical media are grouped by the library into sets of four, and each of these sets is reported as a single storage slot (containing a single virtual media element when tagged as full) to the host computer.

It may also be noted that the move media command of Figure 6 will need to be interpreted differently by such a striped library. For the four drive example, if a move media command issued by the host is to move a virtual

media element at a given address to the virtual drive, the library must move four physical pieces of media, one to each drive. It can further be appreciated that the library response to the write command must include data disassembly and concurrent writes to all four drives.

Mirrored Mode:

In the mirrored mode of operation, media elements and drives are utilized in pairs to simultaneously create two copies of all data written to the library. This will allow data recovery in the event of a failed media element or drive. This can also be useful in that one copy of the data may be permanently resident in the library, and the other copy can be taken off-site. In a mirrored system, of course, the data storage capacity is one half that of the total capacity of the physical media elements, and the data transfer rate is equivalent to an individual drive.

As a representative example, a suitable library according to the present invention may have two drives, and be configured for mirrored mode operation such that the library reports to the host that it contains only one drive, and half the media storage slots than are actually physically present. In this configuration, a move medium command will load a pair of media elements, one into each drive, even though the command will only reference one source and destination address. Once again, the host computer is interfacing with a "virtual" configuration different from the physical configuration of the library. In a similar fashion, the write command will be interpreted as effectively two write commands, one addressed to each drive, so that an identical copy of the data is made on the second drive. In mirrored mode, it may also be advantageous to configure the library such that if a media element is missing, or if the media elements forming the data copy are off site, the corresponding remaining media element or elements are tagged as write protected. This allows read access to the library, but will prevent writes unless the media is available to produce the desired duplicate copy.

Striped with Parity:

This configuration is somewhat similar to a striped media library, but in this case one of the drives is dedicated to the writing of parity data, while the remaining drives function as described above with respect to striping. A suitable system may include five drives, with four for data striping and one for parity. In a manner analogous to the configurations described above, a preferable embodiment allows the host computer to interface with a virtual library configuration comprising a single drive, and one fifth the actual number of media elements as are physically present. In this preferred embodiment, the parity drives writes one block of error detecting and correcting code for each set of blocks written to the four data striping drives. To retain concurrency, the error detecting and correcting block is preferably the same length as each data block of the set. Library response to move medium and write commands is similar to that described above with respect to striped mode, with the complication that an element of parity media must be loaded and written to with internally generated error detection and correction data. Those of skill in the art will recognize that an appropriate choice error correction and block size will allow the complete recovery of data in real time in the event of a failed drive or media element.

It can be appreciated from the foregoing that a preferred media library according to the present invention is user configurable for various combinations of fault tolerance and data transfer performance while retaining an appearance to a host computer system as a conventional media library. The implementation of such fault tolerance and data transfer enhancements requires no modification of application software running on the host system 36. The actual user configuration of the library 10 can occur in several ways. In a preferred media library, the keypad interface 18 allows library 10 configuration via the keypad. Most preferably, the user can scroll through and choose from a list of pre-defined configuration options, or define their own from scratch through the keypad interface. Alternatively, a second SCSI, RS232, or other communication interface can be provided on the library which is dedicated to interfacing with a PC or other computing device which is independent from the host. The PC would run software which would download configuration parameters to the library. As another alternative, a software program which can be run on the host system is used to load the configuration memory 66 with the desired parameters via SCSI communication with the library. This last alternative, however, may require the generation of several separate host configuration control programs which are compatible with different host library control programs being used at different library installations.

The foregoing description details certain preferred embodiments of the present invention and describes the best mode contemplated. It will be appreciated, however, that no matter how detailed the foregoing appears in text, the invention can be practiced in many ways. For example, libraries in accordance with the present invention may use combinations of the several operating modes described above, or may define other "virtual library" configurations which are manipulated by the host system 36. Details of the library status reporting to the host may also vary considerably in its details, while remaining within the spirit and scope of the invention disclosed herein. As is also stated above, it should be noted that the use of particular terminology when describing certain features or aspects of the present invention should not be taken to imply that the broadest reasonable meaning of such terminology is not intended, or that the terminology is being re-defined herein to be restricted to including any specific characteristics of the features or aspects of the invention with which that terminology is associated. The scope of the present invention should therefore be construed in accordance with the appended Claims and any equivalents thereof.

WHAT IS CLAIMED IS:

1. An automated tape cartridge library comprising:

a tape drive;

a plurality of tape cartridges;

a plurality of tape cartridge storage locations;

a moveable carriage adapted to transport tape cartridges from at least one of said tape cartridge storage locations to said tape drive;

a processor coupled to said tape drive, said moveable carriage, and a host computer so as to receive selected commands from said host computer and, in response thereto, send commands to said tape drive and to said moveable carriage, and send information to said host computer; and,

a user configurable memory coupled to said processor and storing data controlling said processor's response to said commands from said host computer so that a user of the library can modify the commands sent to said tape drive and said moveable carriage by said processor, or can modify said information sent to said host computer by said processor in response to said selected commands.

2. The tape cartridge library of Claim 1, wherein said memory comprises non-volatile memory.

3. The tape cartridge library of Claim 1, wherein said library comprises a plurality of tape drives.

4. The tape cartridge library of Claim 3, wherein one of said selected commands comprises an information request for the number of tape drives in said library, and wherein said memory is configured such that in response to said information request said processor informs said host computer that fewer tape drives are physically present in said library than are actually physically present in said library.

5. A method of performing a data storage operation comprising the steps of:

issuing a command from a host computer system to a media element library, said command instructing said media element library to move one media element from a storage location having a first address to one drive having a second address;

in response to said command, loading a plurality of media elements into a respective plurality of drives; and,

writing data to be stored to said plurality of media elements.

6. The method of Claim 5 wherein said writing step comprises the step of writing substantially identical copies of said data to at least two of said plurality of media elements.

7. The method of Claim 5 wherein said step of writing data comprises the steps of breaking said data into portions, and writing different ones of said portions to different ones of said plurality of media elements.

8. The method of Claim 7 wherein each of said plurality of media elements receives an approximately equal amount of data for storage.

9. A method of performing a data retrieval operation comprising the steps of:

issuing a command from a host computer system to a media element library, said command instructing said media element library to move one media element from a storage location having a first address to one drive having a second address;

in response to said command, loading a plurality of media elements into a respective plurality of drives; and,

reading data from said plurality of media elements.

10. The method of Claim 9 wherein said step of reading data comprises the steps of reading portions of said data from each of said plurality of media elements, and assembling said portions of said data into an output data stream.

11. The method of Claim 10 wherein an approximately equal amount of data is read from each of said plurality of media elements.

12. A method of communicating information between a host computer system and a data storage library, said method comprising the steps of:

sending a request for information regarding the number of media element storage locations in said library from the host computer to the library;

sending a response to said request from said library to said host computer system, said response including a number of media element storage locations different from the number of media element storage locations physically present in said library.

13. The method of Claim 12, wherein said response includes a number of media element storage locations which is less than the number of media element storage locations physically present in said library.

14. The method of Claim 12, additionally comprising the steps of:

sending a request for information regarding the number of drives in said library from the host computer to the library;

sending a response to said request from said library to said host computer system, said response including a number of drives different from the number of drives physically present in said library.

15. The method of Claim 14, wherein said response includes a number of media element storage locations which is less than the number of media element storage locations physically present in said library.

16. A media element library comprising:

a plurality of storage locations for removable media elements;

at least one media element drive;

at least one robot for transporting media elements to and from said storage locations and said at least one drive, wherein the number of said storage locations, drives, and robots defines a library configuration; and,

an interface to a host computer, wherein said interface accepts requests for configuration information from said library, and said library responds to said requests with configuration information

different from said configuration defined by said storage locations, said at least one media element drive, and said at least one robot.

17. The media element library of Claim 16, wherein said removable media elements comprise tape cartridges.

18. The media element library of Claim 16, wherein said removable media elements comprise optical disks.

19. The media element library of Claim 16, wherein said interface comprises a SCSI interface.

20. The media element library of Claim 16, additionally comprising a memory, wherein information stored in said memory determines said configuration information contained in said library response.

21. The media element library of Claim 16, wherein said configuration information in said library response defines a second library configuration, and wherein said interface accepts commands from said host computer for manipulating said second library configuration.

22. The media element library of Claim 21, additionally comprising a non-volatile memory, wherein information stored in said non-volatile memory defines said second library configuration.

23. A method of communication between a media element library and a host computer system, said media element library comprising a plurality of media element storage locations, said method comprising the steps of:

sending a first communication to said media library from said host computer;

responding to said first communication from said host computer with a second communication from said media library comprising a plurality of bits, said plurality of bits being in a format which is interpreted by said host computer system as defining a specific number of media element storage locations; and,

interpreting said second communication in said host computer system as a specific number of media element storage locations different from said plurality of media element storage locations physically present in said media element library.

24. A method of communication between a media element library and a host computer system, said media element library comprising a plurality of drives, said method of communication comprising the steps of:

sending a first communication to said media library from said host computer;

responding to said first communication from said host computer with a second communication from said media library comprising a plurality of bits, said plurality of bits being in a format which is interpreted by said host computer system as defining a specific number of drives; and,

interpreting said second communication in said host computer system as a specific number of drives different from said plurality of drives physically present in said media element library.

25. A media element library comprising:

a plurality of media elements;

a plurality of media element storage locations;

at least one media element drive;

at least one robot configured to transport said media elements between said media element storage locations and said at least one media element drive;

means for receiving commands from a host computer system, said commands including requests for information regarding the number of media element storage locations present in said media element library;

means for responding to said commands with information comprising a number of media element storage locations different from the number of said plurality of media element storage locations physically present in said media element library.

26. The media element library of Claim 25, wherein said media elements comprise tape cartridges.

27. The media element library of Claim 25, wherein said media elements comprise optical disks.

28. A method of making a media element library comprising the step of storing data for controlling a response of said media element library to a command from a host computer system in a non-volatile memory in said media element library.

29. The method of Claim 28 additionally comprising the steps of erasing said data in said non-volatile memory and storing different data so as to change said response.

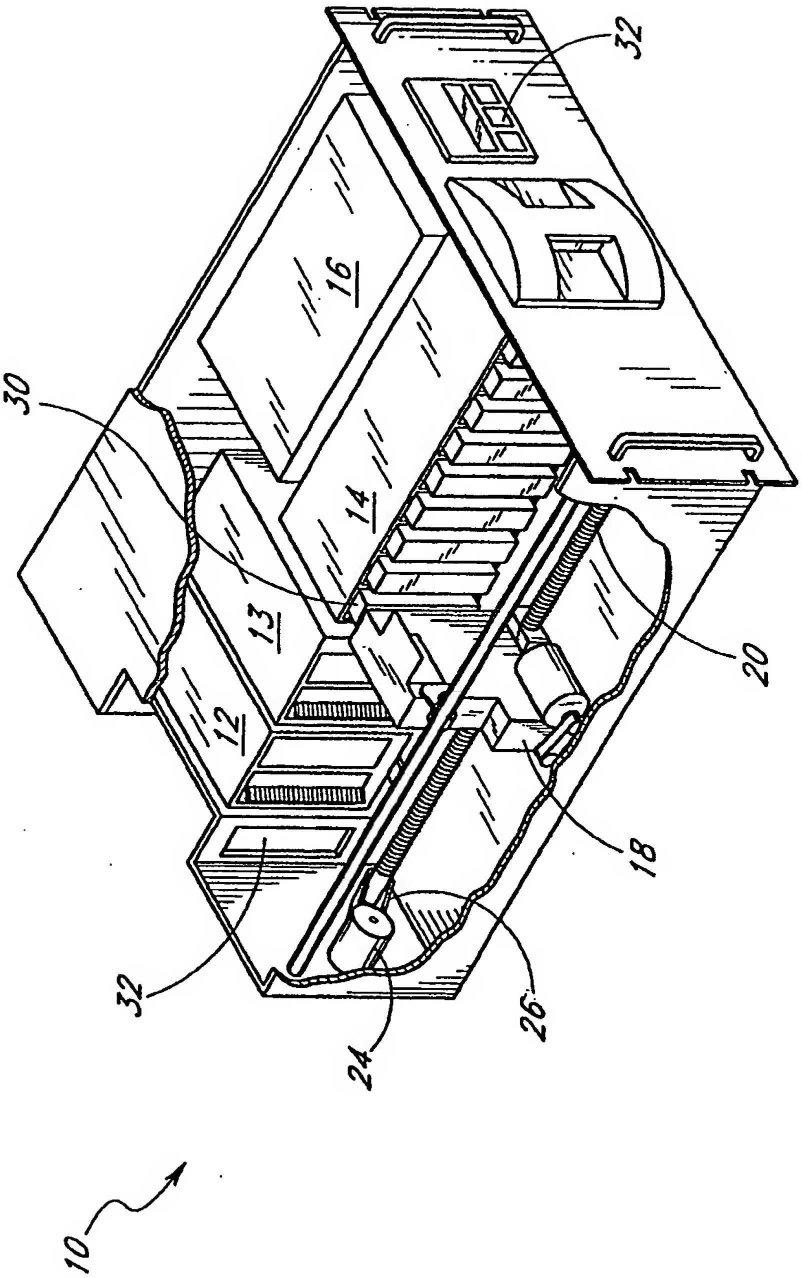


FIG. 1

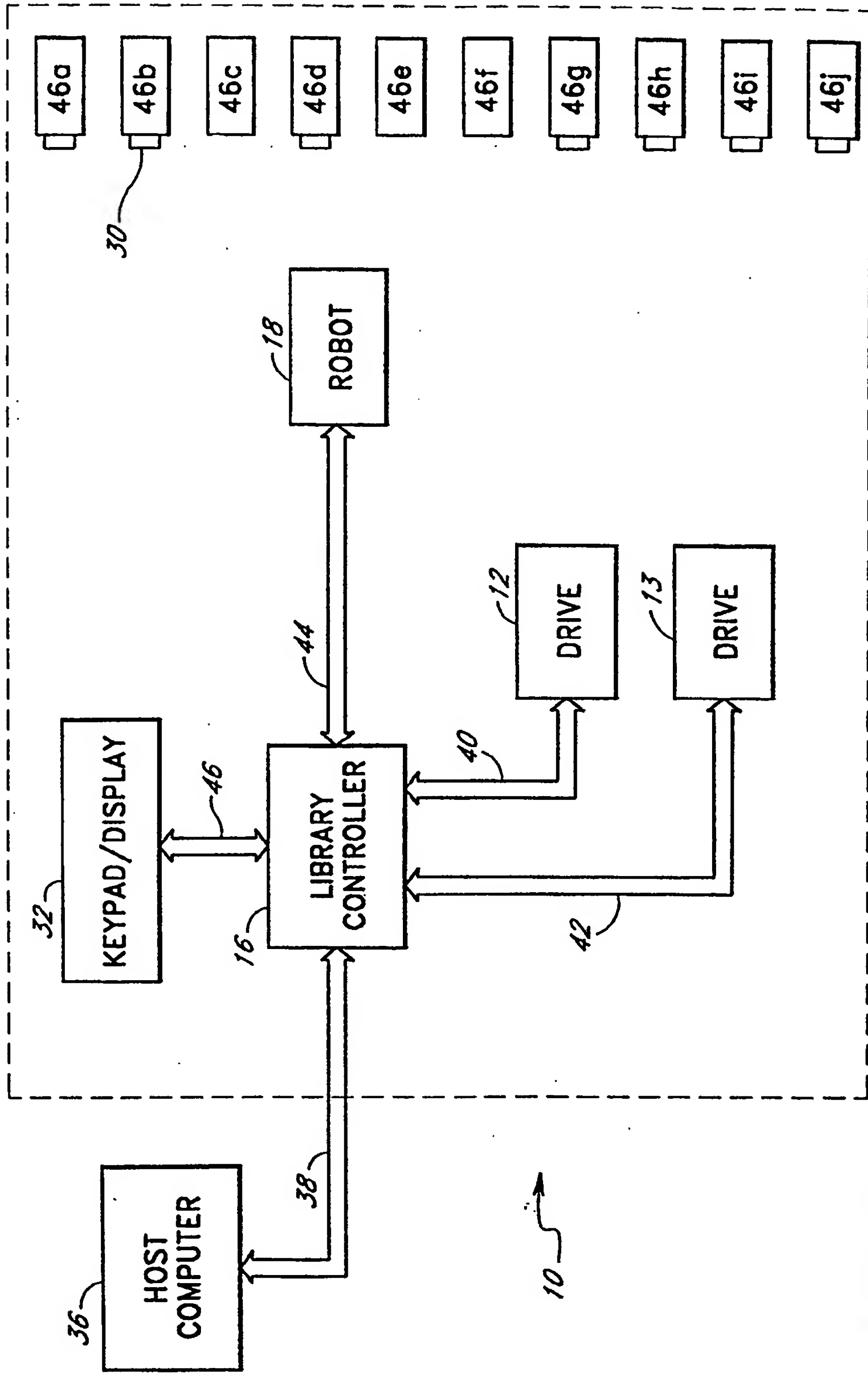

46a
46b
46c
46d
46e
46f
46g
46h
46i
46j
30
ROBOT
18
44
DRIVE
12
DRIVE
13
KEYPAD/DISPLAY
32
46
LIBRARY CONTROLLER
16
40
42
HOST COMPUTER
36
38
10

FIG.2

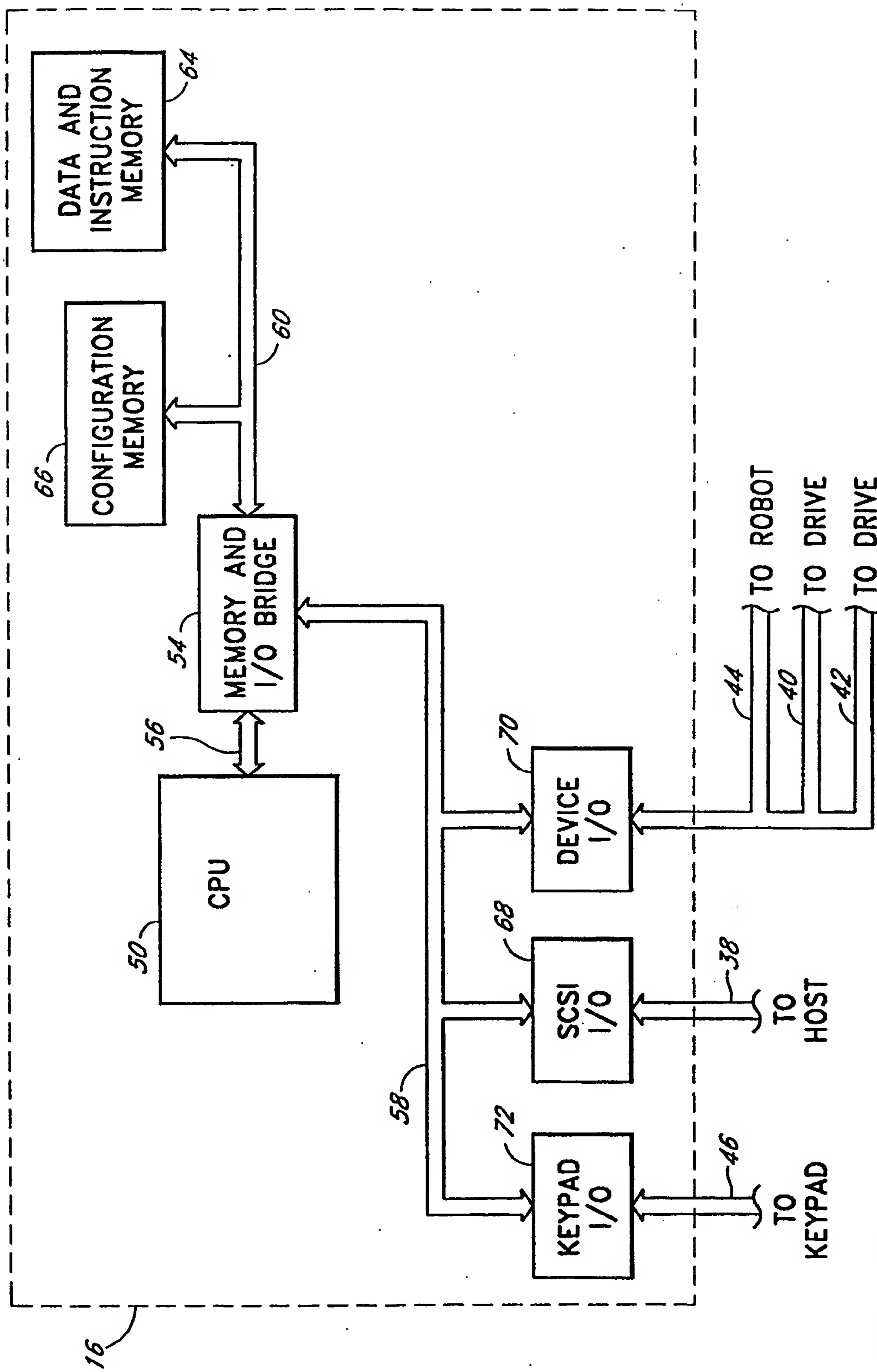

DATA AND INSTRUCTION MEMORY
64
CONFIGURATION MEMORY
66
60
MEMORY AND I/O BRIDGE
54
56
CPU
50
DEVICE I/O
70
SCSI I/O
68
KEYPAD I/O
72
58
44
40
42
TO ROBOT
TO DRIVE
TO DRIVE
38
TO HOST
46
TO KEYPAD
16

FIG.3

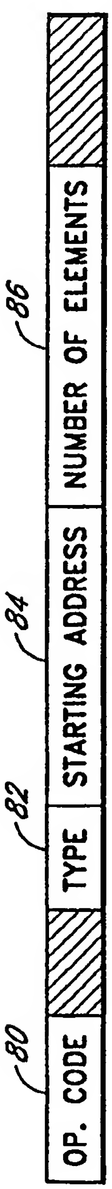


FIG. 4

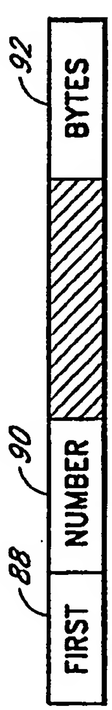


FIG. 5A

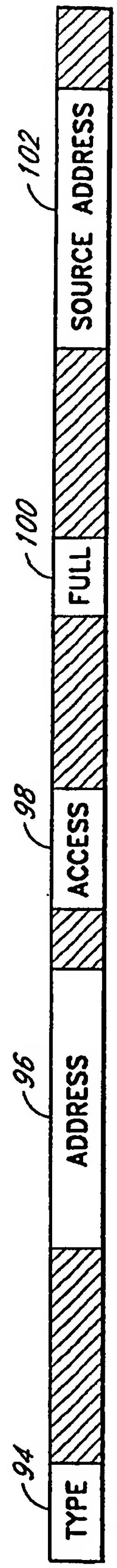


FIG. 5B

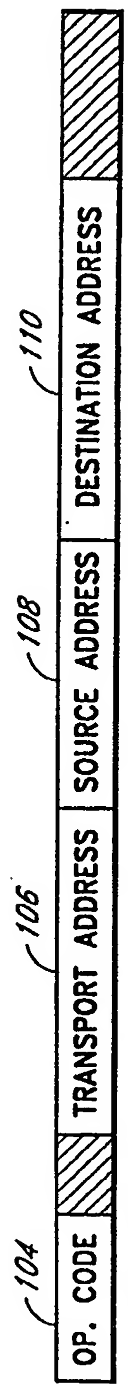


FIG. 6

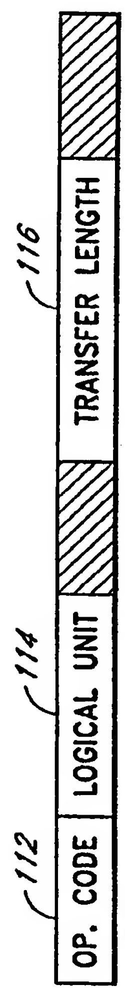


FIG 7

INTERN⁄ ˜IONAL SEARCH REPORT

Inter‹ .al Application No
PCT/US 98/01327

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F3/06

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06F G11B

Documentation searched other than minimumdocumentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 5 386 545 A (GOMBOS JR RICHARD G ET AL) 31 January 1995 see column 4, line 20 - column 10, line 2; figures 1-3	1,16,17, 19 5,9,12, 23-25,28
X A	EP 0 645 693 A (IBM) 29 March 1995 see column 4, line 16 - column 8, line 53; figures 1-4	5-11 1,12,16, 23-25,28
X	"TAPE ARRAY STORAGE SYSTEM" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 36, no. 3, 1 March 1993, pages 369-370, XP000354811 see page 369, paragraph 1 see page 370, paragraph 6	5-11

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publicationdate of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of theinternational search	Date of mailing of the international search report
19 May 1998	29/05/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Moens, R

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 06393 A (ARCADA SOFTWARE INC) 29 February 1996	5,9
A	see abstract see page 8, line 6 - page 10, line 18 see page 14, line 18 - page 15, line 18; figures 1-3B	1-4,6-8, 10-29
A	EP 0 458 567 A (IBM) 27 November 1991 see column 5, line 30 - column 7, line 4 see column 9, line 56 - column 10, line 41 see column 12, line 6 - line 22; figures 1,5,6,8	1-29
A	EP 0 276 967 A (STORAGE TECHNOLOGY CORP) 3 August 1988 see abstract see page 36, line 51 - page 37, line 17 see page 46, line 27 - line 60; figures 1,41A,41B,49	1,5,9, 12,16, 23-25,28
A	EP 0 578 918 A (TANDBERG DATA) 19 January 1994 see abstract see column 3, line 50 - column 4, line 7 see column 5, line 55 - column 6, line 15; figures 2,7	1,2,16, 25,28

1

INTERN TIONAL SEARCH REPORT

In ation on patent family members

Intern al Application No
PCT/US 98/01327

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5386545 A	31-01-95	EP 0664045 A	26-07-95
		JP 8503086 T	02-04-96
		WO 9408337 A	14-04-94
EP 0645693 A	29-03-95	US 5502811 A	26-03-96
		JP 7114448 A	02-05-95
WO 9606393 A	29-02-96	AU 3372095 A	14-03-96
EP 0458567 A	27-11-91	US 5121483 A	09-06-92
		JP 2554407 B	13-11-96
		JP 4228147 A	18-08-92
		US 5386516 A	31-01-95
EP 0276967 A	03-08-88	US 4864511 A	05-09-89
		AT 139360 T	15-06-96
		AU 607489 B	07-03-91
		AU 1541988 A	10-08-88
		CA 1337831 A	26-12-95
		DE 3855355 D	18-07-96
		DE 3855355 T	23-01-97
		ES 2090013 T	16-10-96
		FI 893602 A	27-07-89
		IL 85163 A	25-01-94
		JP 2502139 T	12-07-90
		WO 8805593 A	28-07-88
		US 4928245 A	22-05-90
		US 4932826 A	12-06-90
		ZA 8800493 A	12-07-88
EP 0578918 A	19-01-94	US 5357495 A	18-10-94
		JP 2566725 B	25-12-96
		JP 6124531 A	06-05-94